

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
25.08.2004 Bulletin 2004/35

(51) Int Cl.7: **G10L 21/02**

(21) Application number: **04003675.8**

(22) Date of filing: **18.02.2004**

<div>(84) Designated Contracting States: <b>AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PT RO SE SI SK TR</b> Designated Extension States: <b>AL LT LV MK</b></div> <div>(30) Priority: <b>21.02.2003 US 449511 P</b> <b>10.04.2003 US 410736</b> <b>16.10.2003 US 688802</b></div> <div>(71) Applicant: <b>Harman Becker Automotive Systems-Wavemakers, Inc.</b> <b>Vancouver, British Columbia, V6B 2K4 (CA)</b></div>	<div>(72) Inventors: • <b>Hetherington, Phil</b> <b>Vancouver B.C. V6M 269 (CA)</b> • <b>Li, Xueman</b> <b>Barneley B.C. V5B 4T6 (CA)</b> • <b>Zakarauskas, Pierre</b> <b>Vancouver B.C. V52 4B2 (CA)</b></div> <div>(74) Representative: <b>Grünecker, Kinkeldey, Stockmair &amp; Schwanhäusser Anwaltssozietät</b> <b>Maximilianstrasse 58</b> <b>80538 München (DE)</b></div>
---	--

(54) **System for suppressing wind noise**

(57) A voice enhancement logic improves the perceptual quality of a processed voice. The voice enhancement system includes a noise detector and a noise attenuator. The noise detector detects a wind buffet and a continuous noise by modeling the wind buffet. The noise attenuator dampens the wind buffet to improve the intelligibility of an unvoiced, a fully voiced, or a mixed voice segment.

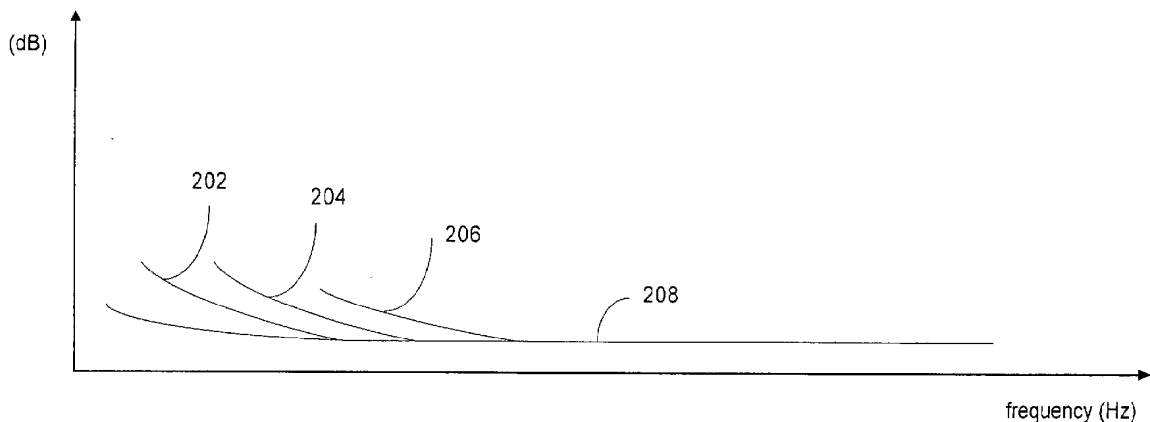


FIGURE 2

## Description

[0001] This application is a continuation in-part of U. S. Application No. 10/410,736, "Method and Apparatus for Suppressing Wind Noise," filed April 10, 2003. The disclosure of the above application is incorporated herein by reference.

## BACKGROUND OF THE INVENTION

### 1. Technical Field.

[0002] This invention relates to acoustics, and more particularly, to a system that enhances the perceptual quality of a processed voice.

### 2. Related Art.

[0003] Many hands-free communication devices acquire, assimilate, and transfer a voice signal. Voice signals pass from one system to another through a communication medium. In some systems, including some used in vehicles, the clarity of the voice signal does not depend on the quality of the communication system or the quality of the communication medium. When noise occurs near a source or a receiver, distortion garbles the voice signal, destroys information, and in some instances, masks the voice signal so that it is not recognized by a listener.

[0004] Noise, which may be annoying, distracting, or results in a loss of information, may come from many sources. Within a vehicle, noise may be created by the engine, the road, the tires, or by the movement of air. A natural or artificial movement of air may be heard across a broad frequency range. Continuous fluctuations in amplitude and frequency may make wind noise difficult to overcome and degrade the intelligibility of a voice signal.

[0005] Many systems attempt to counteract the effects of wind noise. Some systems rely on a variety of sound-suppressing and dampening materials throughout an interior to ensure a quiet and comfortable environment. Other systems attempt to average out varying wind-induced pressures that press against a receiver. These noise reducers may take many shapes to filter out selected pressures making them difficult to design to the many interiors of a vehicle. Another problem with some speech enhancement systems is that of detecting wind noise in a background of a continuous noise. Yet another problem with some speech enhancement systems is that they do not easily adapt to other communication systems that are susceptible to wind noise.

[0006] Therefore there is a need for a system that counteracts wind noise across a varying frequency range.

## SUMMARY

[0007] A voice enhancement logic improves the per-

ceptual quality of a processed voice. The system learns, encodes, and then dampens the noise associated with the movement of air from an input signal. The system includes a noise detector and a noise attenuator. The noise detector detects a wind buffet by modeling. The noise attenuator then dampens the wind buffet.

Alternative voice enhancement logic includes time frequency transform logic, a background noise estimator, a wind noise detector, and a wind noise attenuator. The time frequency transform logic converts a time varying input signal into a frequency domain output signal. The background noise estimator measures the continuous noise that may accompany the input signal. The wind noise detector automatically identifies and models a wind buffet, which may then be dampened by the wind noise attenuator.

[0008] Other systems, methods, features and advantages of the invention will be, or will become, apparent to one with skill in the art upon examination of the following figures and detailed description. It is intended that all such additional systems, methods, features and advantages be included within this description, be within the scope of the invention, and be protected by the following claims.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The invention can be better understood with reference to the following drawings and description. The components in the figures are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention. Moreover, in the figures, like referenced numerals designate corresponding parts throughout the different views.

[0010] Figure 1 is a partial block diagram of voice enhancement logic.

[0011] Figure 2 is noise that may be associated with wind and other sources in the frequency domain.

[0012] Figure 3 is a signal-to-noise ratio of the noise that may be associated with wind and other sources in the frequency domain.

[0013] Figure 4 is a block diagram of the voice enhancement logic of Figure 1.

[0014] Figure 5 is a pre-processing system coupled to the voice enhancement logic of Figure 1.

[0015] Figure 6 is an alternative pre-processing system coupled to the voice enhancement logic of Figure 1.

[0016] Figure 7 is a block diagram of an alternative voice enhancement system.

[0017] Figure 8 is noise that may be associated with wind and other sources in the frequency domain.

[0018] Figure 9 is a graph of a wind buffet masking a portion of a voice signal.

[0019] Figure 10 is a graph of a processed and reconstructed voice signal.

[0020] Figure 11 is a flow diagram of a voice enhancement.

[0021] Figure 12 is a partial sequence diagram of a

voice enhancement.

**[0022]** Figure 13 is a partial sequence diagram of a voice enhancement.

**[0023]** Figure 14 is a block diagram of voice enhancement logic within a vehicle.

**[0024]** Figure 15 is a block diagram of voice enhancement logic interfaced to an audio system and/or a communication system.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

**[0025]** A voice enhancement logic improves the perceptual quality of a processed voice. The logic may automatically learn and encode the shape and form of the noise associated with the movement of air in a real or a delayed time. By tracking selected attributes, the logic may eliminate or dampen wind noise using a limited memory that temporarily stores the selected attributes of the noise. Alternatively, the logic may also dampen a continuous noise and/or the "musical noise," squeaks, squawks, chirps, clicks, drips, pops, low frequency tones, or other sound artifacts that may be generated by some voice enhancement systems.

**[0026]** Figure 1 is a partial block diagram of the voice enhancement logic 100. The voice enhancement logic may encompass hardware or software that is capable of running on one or more processors in conjunction with one or more operating systems. The highly portable logic includes a wind noise detector 102 and a noise attenuator 104.

**[0027]** In Figure 1 the wind noise detector 102 may identify and model a noise associated with wind flow from the properties of air. While wind noise occurs naturally or may be artificially generated over a broad frequency range, the wind noise detector 102 is configured to detect and model the wind noise that is perceived by the ear. The wind noise detector receives incoming sound, that in the short term spectra, may be classified into three broad categories: (1) unvoiced, which exhibits noise-like characteristics that includes the noise associated with wind, i.e., it may have some spectral shape but no harmonic or formant structure; (2) fully voiced, which exhibits a regular harmonic structure, or peaks at pitch harmonics weighted by the spectral envelope that may describe the formant structure, and (3) mixed voice, which exhibits a mixture of the above two categories, some parts containing noise-like segments, the rest exhibiting a regular harmonic structure and/or a formant structure.

**[0028]** The wind noise detector 102 may separate the noise-like segments from the remaining signal in a real or in a delayed time no matter how complex or how loud an incoming segment may be. The separated noise-like segments are analyzed to detect the occurrence of wind noise, and in some instances, the presence of a continuous underlying noise. When wind noise is detected, the spectrum is modeled, and the model is retained in a

memory. While the wind noise detector 102 may store an entire model of a wind noise signal, it also may store selected attributes in a memory.

**[0029]** To overcome the effects of wind noise, and in some instances, the underlying continuous noise that may include ambient noise, the noise attenuator 104 substantially removes or dampens the wind noise and/or the continuous noise from the unvoiced and mixed voice signals. The voice enhancement logic 100 encompasses any system that substantially removes or dampens wind noise. Examples of systems that may dampen or remove wind noise include systems that use a signal and a noise estimate such as (1) systems which use a neural network mapping of a noisy signal and an estimate of the noise to a noise-reduced signal, (2) systems which subtract the noise estimate from a noisy-signal, (3) systems that use the noisy signal and the noise estimate to select a noise-reduced signal from a codebook, (4) systems that in any other way use the noisy signal and the noise estimate to create a noise-reduced signal based on a reconstruction of the masked signal. These systems may attenuate wind noise, and in some instances, attenuate the continuous noise that may be part of the short-term spectra. The noise attenuator 104 may also interface or include an optional residual attenuator 106 that removes or dampens artifacts that may result in the processed signal. The residual attenuator 106 may remove the "musical noise," squeaks, squawks, chirps, clicks, drips, pops, low frequency tones, or other sound artifacts.

**[0030]** Figure 2 illustrates exemplary noise associated with three wind flows. The wind buffets 202, 204, and 206, which are the events of wind striking a detector, vary by their level of severity or amplitude. The amplitudes reflect the relative differences in power or intensity between the fluctuations of air pressure received across an input area of a receiver or a detector. The line underlying the wind buffets illustrates the continuous noise 208 that is also sensed by the receiver or detector. In a vehicle, wind buffets may represent the natural flow of air through a window, through an open top of a convertible, through an inlet, or the artificial movement of air caused by a fan or a heating, ventilating, and/or air conditioning system (HVAC). The continuous noise may represent an ambient noise or a noise associated with an engine, a powertrain, a road, tires, or other sounds.

**[0031]** In the time and frequency spectral domain, the continuous noise 208 and a wind buffet 202 may be curvilinear. The continuous noise and wind buffet may appear to be formed or characterized by the curved lines shown in Figure 2. However, when the signal strength (in decibels) of the wind buffet (e.g.,  $\sigma_{WB}$ ) is related to the signal strength of a continuous noise (e.g.,  $\sigma_{CN}$ ) in the signal-to-noise ratio (SNR) domain, the wind buffet 202 may be characterized by a linear function with a vertical dimension corresponding to decibels and a horizontal dimension corresponding to frequency. This relation may be expressed as:

$$\text{SNR} = \sigma_{\text{WB}} - \sigma_{\text{CN}} \quad (\text{Equation 1})$$

Any method may approximate the linearity of a wind buffet. In the signal-to-noise domain, an offset or y-intercept 302 and an x-intercept or pivot point may characterize the linear model 302. Alternatively, an x or y-coordinate and a slope may model the wind buffet. In Figure 3, the linear model 302 descends in a negative slope.

**[0032]** Figure 4 is a block diagram of an example wind noise detector 102 that may receive or detect an unvoiced, fully voiced, or a mixed voice input signal. A received or detected signal is digitized at a predetermined frequency. To assure a good quality voice, the voice signal is converted to a pulse-code-modulated (PCM) signal by an analog-to-digital converter 402 (ADC) having any common sample rate. A smooth window 404 is applied to a block of data to obtain the windowed signal. The complex spectrum for the windowed signal may be obtained by means of a fast Fourier transform (FFT) 406 that separates the digitized signals into frequency bins, with each bin identifying an amplitude and phase across a small frequency range. Each frequency bin may then be converted into the power-spectral domain 408 and logarithmic domain 410 to develop a wind buffet and continuous noise estimate. As more windows of sound are processed, the wind noise detector 102 may derive average noise estimates. A time-smoothed or weighted average may be used to estimate the wind buffet and continuous noise estimates for each frequency bin.

**[0033]** To detect a wind buffet, a line may be fitted to a selected portion of the low frequency spectrum in the SNR domain. Through a regression, a best-fit line may measure the severity of the wind noise within a given block of data. A high correlation between the best-fit line and the low frequency spectrum may identify a wind buffet. Whether or not a high correlation exists, may depend on a desired clarity of a processed voice and the variations in frequency and amplitude of the wind buffet. Alternatively, a wind buffet may be identified when an offset or y-intercept of the best-fit line exceeds a predetermined threshold (e.g., > 3 dB).

**[0034]** To limit a masking of voice, the fitting of the line to a suspected wind buffet signal may be constrained by rules. Exemplary rules may prevent a calculated offset, slope, or coordinate point in a wind buffet model from exceeding an average value. Another rule may prevent the wind noise detector 102 from applying a calculated wind buffet correction when a vowel or another harmonic structure is detected. A harmonic may be identified by its narrow width and its sharp peak, or in conjunction with a voice or a pitch detector. If a vowel or another harmonic structure is detected, the wind noise detector may limit the wind buffet correction to values less than or equal to average values. An additional rule may allow the average wind buffet model or its attributes to be updated only during unvoiced segments.

If a voiced or a mixed voice segment is detected, the average wind buffet model or its attributes are not updated under this rule. If no voice is detected, the wind buffet model or each attribute may be updated through any means, such as through a weighted average or a leaky integrator. Many other rules may also be applied to the model. The rules may provide a substantially good linear fit to a suspected wind buffet without masking a voice segment.

**[0035]** To overcome the effects of wind noise, a wind noise attenuator 104 may substantially remove or dampen the wind buffet from the noisy spectrum by any method. One method may add the wind buffet model to a recorded or modeled continuous noise. In the power spectrum, the modeled noise may then be subtracted from the unmodified spectrum. If an underlying peak or valley 902 is masked by a wind buffet 202 as shown in Figure 9 or masked by a continuous noise, a conventional or modified interpolation method may be used to reconstruct the peak and/or valley as shown in Figure 10. A linear or step-wise interpolator may be used to reconstruct the missing part of the signal. An inverse FFT may then be used to convert the signal power to the time domain, which provides a reconstructed voice signal.

**[0036]** To minimize the "music noise," squeaks, squawks, chirps, clicks, drips, pops, low frequency tones, or other sound artifacts that may be generated in the low frequency range by some wind noise attenuators, an optional residual attenuator 106 (shown in Figure 1) may also condition the voice signal before it is converted to the time domain. The residual attenuator 106 may track the power spectrum within a low frequency range (e.g., less than about 400 Hz). When a large increase in signal power is detected an improvement may be obtained by limiting or dampening the transmitted power in the low frequency range to a predetermined or calculated threshold. A calculated threshold may be equal to, or based on, the average spectral power of that same low frequency range at an earlier period in time.

**[0037]** Further improvements to voice quality may be achieved by pre-conditioning the input signal before the wind noise detector processes it. One pre-processing system may exploit the lag time that a signal may arrive at different detectors that are positioned apart as shown in Figure 5. If multiple detectors or microphones 502 are used that convert sound into an electric signal, the pre-processing system may include control logic 504 that automatically selects the microphone 502 and channel that senses the least amount of noise. When another microphone 502 is selected, the electric signal may be combined with the previously generated signal before being processed by the wind noise detector 102.

**[0038]** Alternatively, multiple wind noise detectors 102 may be used to analyze the input of each of the microphones 502 as shown in Figure 6. Spectral wind buffet estimates may be made on each of the channels. A mixing of one or more channels may occur by switching between the outputs of the microphones 502. The

signals may be evaluated and selected on a frequency-by-frequency basis until the frequency of the pivot point 304 (shown in Figure 3) is reached. Alternatively, control logic 602 may combine the output signals of multiple wind noise detectors 102 at a specific frequency or frequency range through a weighting function. When the frequency of the pivot point is exceeded, the process may continue or a standard adaptive beam forming method may be used.

**[0039]** Figure 7 is alternative voice enhancement logic 700 that also improves the perceptual quality of a processed voice. The enhancement is accomplished by time-frequency transform logic 702 that digitizes and converts a time varying signal to the frequency domain. A background noise estimator 704 measures the continuous or ambient noise that occurs near a sound source or the receiver. The background noise estimator 704 may comprise a power detector that averages the acoustic power in each frequency bin. To prevent biased noise estimations at transients, a transient detector 706 disables the noise estimation process during abnormal or unpredictable increases in power. In Figure 7, the transient detector 706 disables the background noise estimator 704 when an instantaneous background noise  $B(f, i)$  exceeds an average background noise  $B(f)_{Ave}$  by more than a selected decibel level 'c.' This relationship may be expressed as:

$$B(f, i) > B(f)_{Ave} + c \quad (\text{Equation 2})$$

**[0040]** To detect a wind buffet, a wind noise detector 708 may fit a line to a selected portion of the spectrum in the SNR domain. Through a regression, a best-fit line may model the severity of the wind noise 202, as shown in Figure 8. To limit any masking of voice, the fitting of the line to a suspected wind buffet may be constrained by the rules described above. A wind buffet may be identified when the offset or y-intercept of the line exceeds a predetermined threshold or when there is a high correlation between a fitted line and the noise associated with a wind buffet. Whether or not a high correlation exists, may depend on a desired clarity of a processed voice and the variations in frequency and amplitude of the wind buffet.

**[0041]** Alternatively, a wind buffet may be identified by the analysis of time varying spectral characteristics of the input signal that may be graphically displayed on a spectrograph. A spectrograph may produce a two dimensional pattern called a spectrogram in which the vertical dimensions correspond to frequency and the horizontal dimensions correspond to time.

**[0042]** A signal discriminator 710 may mark the voice and noise of the spectrum in real or delayed time. Any method may be used to distinguish voice from noise. In Figure 7, voiced signals may be identified by (1) the narrow widths of their bands or peaks; (2) the resonant

structure that may be harmonically related; (3) the resonances or broad peaks that correspond to formant frequencies; (4) characteristics that change relatively slowly with time; (5) their durations; and when multiple detectors or microphones are used, (6) the correlation of the output signals of the detectors or microphones.

**[0043]** To overcome the effects of wind noise, a wind noise attenuator 712 may dampen or substantially remove the wind buffet from the noisy spectrum by any method. One method may add the substantially linear wind buffet model to a recorded or modeled continuous noise. In the power spectrum, the modeled noise may then be removed from the unmodified spectrum by the means described above. If an underlying peak or valley 902 is masked by a wind buffet 202 as shown in Figure 9 or masked by a continuous noise, a conventional or modified interpolation method may be used to reconstruct the peak and/or valley as shown in Figure 10. A linear or step-wise interpolator may be used to reconstruct the missing part of the signal. A time series synthesizer may then be used to convert the signal power to the time domain, which provides a reconstructed voice signal.

**[0044]** To minimize the "musical noise," squeaks, squawks, chirps, clicks, drips, pops, low frequency tones, or other sound artifacts that may be generated in the low frequency range by some wind noise attenuators, an optional residual attenuator 714 may also be used. The residual attenuator 714 may track the power spectrum within a low frequency range. When a large increase in signal power is detected an improvement may be obtained by limiting the transmitted power in the low frequency range to a predetermined or calculated threshold. A calculated threshold may be equal to or based on the average spectral power of that same low frequency range at a period earlier in time.

**[0045]** Figure 11 is a flow diagram of a voice enhancement that removes some wind buffets and continuous noise to enhance the perceptual quality of a processed voice. At act 1102 a received or detected signal is digitized at a predetermined frequency. To assure a good quality voice, the voice signal may be converted to a PCM signal by an ADC. At act 1104 a complex spectrum for the windowed signal may be obtained by means of an FFT that separates the digitized signals into frequency bins, with each bin identifying an amplitude and a phase across a small frequency range.

**[0046]** At act 1106, a continuous or ambient noise is measured. The background noise estimate may comprise an average of the acoustic power in each frequency bin. To prevent biased noise estimations at transients, the noise estimation process may be disabled during abnormal or unpredictable increases in power at act 1108. The transient detection act 1108 disables the background noise estimate when an instantaneous background noise exceeds an average background noise by more than a predetermined decibel level.

**[0047]** At act 1110, a wind buffet may be detected

when the offset exceeds a predetermined threshold (e.g., a threshold  $> 3$  dB) or when a high correlation exists between a best-fit line and the low frequency spectrum. Alternatively, a wind buffet may be identified by the analysis of time varying spectral characteristics of the input signal. When a line fitting detection method is used, the fitting of the line to the suspected wind buffet signal may be constrained by some optional acts. Exemplary optional acts may prevent a calculated offset, slope, or coordinate point in a wind buffet model from exceeding an average value. Another optional act may prevent the wind noise detection method from applying a calculated wind buffet correction when a vowel or another harmonic structure is detected. If a vowel or another harmonic structure is detected, the wind noise detection method may limit the wind buffet correction to values less than or equal to average values. An additional optional act may allow the average wind buffet model or attributes to be updated only during unvoiced segments. If a voiced or mixed voice segment is detected, the average wind buffet model or attributes are not updated under this act. If no voice is detected, the wind buffet model or each attribute may be updated through many means, such as through a weighted average or a leaky integrator. Many other optional acts may also be applied to the model.

**[0048]** At act 1112, a signal analysis may discriminate or mark the voice signal from the noise-like segments. Voiced signals may be identified by, for example, (1) the narrow widths of their bands or peaks; (2) the resonant structure that may be harmonically related; (3) their harmonics that correspond to formant frequencies; (4) characteristics that change relatively slowly with time; (5) their durations; and when multiple detectors or microphones are used, (6) the correlation of the output signals of the detectors or microphones.

**[0049]** To overcome the effects of wind noise, a wind noise is substantially removed or dampened from the noisy spectrum by any act. One exemplary act 1114 adds the substantially linear wind buffet model to a recorded or modeled continuous noise. In the power spectrum, the modeled noise may then be substantially removed from the unmodified spectrum by the methods and systems described above. If an underlying peak or valley 902 is masked by a wind buffet 202 as shown in Figure 9 or masked by a continuous noise, a conventional or modified interpolation method may be used to reconstruct the peak and/or valley at act 1116. A time series synthesis may then be used to convert the signal power to the time domain at act 1120, which provides a reconstructed voice signal.

**[0050]** To minimize the "musical noise," squeaks, squawks, chirps, clicks, drips, pops, low frequency tones, or other sound artifacts that may be generated in the low frequency range by some wind noise processes, a residual attenuation method may also be performed before the signal is converted back to the time domain. An optional residual attenuation method 1118 may track

the power spectrum within a low frequency range. When a large increase in signal power is detected an improvement may be obtained by limiting the transmitted power in the low frequency range to a predetermined or calculated threshold. A calculated threshold may be equal to or based on the average spectral power of that same low frequency range at a period earlier in time.

**[0051]** Figures 12 and 13 are partial sequence diagrams of a voice enhancement. Like the method shown in Figure 11, the sequence diagrams may be encoded in a signal bearing medium, a computer readable medium such as a memory, programmed within a device such as one or more integrated circuits, or processed by a controller or a computer. If the methods are performed by software, the software may reside in a memory resident to or interfaced to the wind noise detector 102, a communication interface, or any other type of non-volatile or volatile memory interfaced or resident to the voice enhancement logic 100 or 700. The memory may include an ordered listing of executable instructions for implementing logical functions. A logical function may be implemented through digital circuitry, through source code, through analog circuitry, or through an analog source such through an analog electrical, audio, or video signal. The software may be embodied in any computer-readable or signal-bearing medium, for use by, or in connection with an instruction executable system, apparatus, or device. Such a system may include a computer-based system, a processor-containing system, or another system that may selectively fetch instructions from an instruction executable system, apparatus, or device that may also execute instructions.

**[0052]** A "computer-readable medium," "machine-readable medium," "propagated-signal" medium, and/or "signal-bearing medium" may comprise any means that contains, stores, communicates, propagates, or transports software for use by or in connection with an instruction executable system, apparatus, or device. The machine-readable medium may selectively be, but not limited to, an electronic, magnetic, optical, electromagnetic, infrared, or semiconductor system, apparatus, device, or propagation medium. A non-exhaustive list of examples of a machine-readable medium would include: an electrical connection "electronic" having one or more wires, a portable magnetic or optical disk, a volatile memory such as a Random Access Memory "RAM" (electronic), a Read-Only Memory "ROM" (electronic), an Erasable Programmable Read-Only Memory (EPROM or Flash memory) (electronic), or an optical fiber (optical). A machine-readable medium may also include a tangible medium upon which software is printed, as the software may be electronically stored as an image or in another format (e.g., through an optical scan), then compiled, and/or interpreted or otherwise processed. The processed medium may then be stored in a computer and/or machine memory.

**[0053]** As shown in the first sequence of Figure 12, a time series signal may be digitized and smoothed by a

Hanning window to provide an accurate estimation of a fully voiced, a mixed voice, or an unvoiced segment. The complex spectrum for the windowed signal is obtained by means of an FFT that separates the digitized signals into frequency bins, with each bin identifying an amplitude across a small frequency range.

**[0054]** In the second sequence, an averaging of the acoustic power in each frequency bin during unvoiced segments derives the background noise estimate. To prevent biased noise estimates, noise estimates may not occur when abnormal or unpredictable power fluctuations are detected.

**[0055]** In the third sequence, the unmodified spectrum is digitized, smoothed by a window, and transformed into the complex spectrum by an FFT. The unmodified spectrum exhibits portions containing noise-like segments and other portions exhibiting a regular harmonic structure.

**[0056]** In the fourth sequence, a sound segment is fitted to separate lines to model the severity of the wind and continuous noise. To provide a more complete explanation, an unvoiced, fully voiced, and mixed voiced sample are shown. The frequency bins in each sample were converted into the power-spectral domain and logarithmic domain to develop a wind buffet and continuous noise estimate. As more windows are processed, the average wind noise and continuous noise estimates are derived.

**[0057]** To detect a wind buffet, a line is fitted to a selected portion of the signal in the SNR domain. Through a regression, best-fit lines model the severity of the wind noise in each illustration. A high correlation between one best-fit line and the low frequency spectrum may identify a wind buffet. Alternatively, a y-intercept that exceeds a predetermined threshold may also identify a wind buffet. To limit the masking of voice, the fitting of the line to a suspected wind buffet signal may be constrained by the rules described above.

**[0058]** To overcome the effects of wind noise, the modeled noise may be dampened in the unmodified spectrum. In Figure 13, the dampening of the wind buffets and continuous noise from the unvoiced and mixed voiced sample are shown in the fifth sequence. An inverse FFT that converts the signal power to the time domain provides the reconstructed voice signal.

**[0059]** From the foregoing descriptions it should be apparent that the above-described systems may condition signals received from only one microphone or detector. It should also be apparent, that many combinations of systems may be used to identify and track wind buffets. Besides the fitting of a line to a suspected wind buffet, a system may (1) detect the peaks in the spectra having a SNR greater than a predetermined threshold; (2) identify the peaks having a width greater than a predetermined threshold; (3) identify peaks that lack a harmonic relationships; (4) compare peaks with previous voiced spectra; and (5) compare signals detected from different microphones before differentiating the wind

buffet segments, other noise like segments, and regular harmonic structures. One or more of the systems described above may also be used in alternative voice enhancement logic.

**[0060]** Other alternative voice enhancement systems include combinations of the structure and functions described above. These voice enhancement systems are formed from any combination of structure and function described above or illustrated within the attached figures. The logic may be implemented in software or hardware. The term "logic" is intended to broadly encompass a hardware device or circuit, software, or a combination. The hardware may include a processor or a controller having volatile and/or non-volatile memory and may also include interfaces to peripheral devices through wireless and/or hardwire mediums.

**[0061]** The voice enhancement logic is easily adaptable to any technology or devices. Some voice enhancement systems or components interface or couple vehicles as shown in Figure 14, instruments that convert voice and other sounds into a form that may be transmitted to remote locations, such as landline and wireless telephones and audio equipment as shown in Figure 15, and other communication systems that may be susceptible to wind noise.

**[0062]** The voice enhancement logic improves the perceptual quality of a processed voice. The logic may automatically learn and encode the shape and form of the noise associated with the movement of air in a real or a delayed time. By tracking selected attributes, the logic may eliminate or dampen wind noise using a limited memory that temporarily or permanently stores selected attributes of the wind noise. The voice enhancement logic may also dampen a continuous noise and/or the squeaks, squawks, chirps, clicks, drips, pops, low frequency tones, or other sound artifacts that may be generated within some voice enhancement systems and may reconstruct voice when needed.

**[0063]** While various embodiments of the invention have been described, it will be apparent to those of ordinary skill in the art that many more embodiments and implementations are possible within the scope of the invention. Accordingly, the invention is not to be restricted except in light of the attached claims and their equivalents.

## Claims

**1.** A system for suppressing wind noise from a voiced or unvoiced signal, comprising:

a noise detector that detects and models a wind buffet from an input signal; and  
a noise attenuator electrically connected to the noise detector to substantially remove the wind buffet from the input signal.

2. The system for suppressing wind noise of claim 1 where the noise detector models a line to a portion of the input signal.
3. The system of claim 2 where the noise detector is configured to fit a line to a portion of the input signal in a SNR domain.
4. The system of claim 1 where the noise detector is configured to model the wind buffet by calculating a signal offset.
5. The system of claim 1 where the noise detector is configured to prevent the attributes of the modeled wind buffet from exceeding their respective average values.
6. The system of claim 1 where the noise detector is configured to limit a wind buffet correction when a vowel or a harmonic like structure is detected.
7. The system of claim 1 where the noise detector is configured to derive an average wind buffet model, and the average wind buffet model is not updated when a voiced or a mixed voice signal is detected.
8. The system of claim 1 where the noise detector is configured to derive an average wind buffet model that is derived by a weighted average of other modeled signals analyzed earlier in time.
9. The system of claim 1 where the noise attenuator is configured to substantially remove the wind buffet and a continuous noise from the input signal.
10. The system of claim 1 further comprising a residual attenuator electrically coupled to the noise detector and the noise attenuator to dampen signal power in a low frequency range when a large increase in a signal power is detected in the low frequency range.
11. The system of claim 1 further including an input device electrically coupled to the noise detector, the input device configured to convert sound waves into analog signals.
12. The system of claim 1 further including a pre-processing system coupled to the noise detector, the pre-processing system configured to pre-condition the input signal before the wind noise detector processes it.
13. The system of claim 12 where the pre-processing system comprises first and second microphones spaced apart and configured to exploit a lag time of a signal that may arrive at the different detectors.
14. The system of claim 13 further comprising control logic that automatically selects a microphone and a channel that senses the least amount of noise in the input signal.
15. The system of claim 13 further comprising a second noise detector coupled to the noise detector and the first microphone.
16. A system for detecting wind noise from a voiced and unvoiced signal, comprising:
  - a time frequency transform logic that converts a time varying input signal into the frequency domain;
  - a background noise estimator coupled to the time frequency transform logic, the background noise estimator configured to measure the continuous noise that occurs near a receiver; and
  - a wind noise detector coupled to the background noise estimator, the wind noise detector configured to automatically identify and model a noise associated with wind.
17. The system of claim 16 further comprising a transient detector configured to disable the background noise estimator when a transient signal is detected.
18. The system of claim 16 where the wind noise detector is configured to derive a correlation between the line and a portion of the input signal.
19. The system of claim 16 further comprising a signal discriminator coupled to the wind noise detector, the signal discriminator configured to mark the voice and the noise segments of the input signal.
20. The system of claim 16 further comprising a wind noise attenuator coupled to the wind noise detector, the wind noise attenuator configured to reduce the noise associated with the wind that is sensed by the receiver.
21. The system of claim 16 where the noise attenuator is configured to substantially remove the noise associated with the wind from the input signal.
22. The system of claim 16 further comprising a residual attenuator coupled to the background noise estimator operable to dampen signal power in a low frequency range when a large increase in signal power is detected in the low frequency range.
23. A system for suppressing wind noise from a voiced or unvoiced signal, comprising:
  - a time frequency transform logic that converts a time varying input signal into the frequency domain;



- a background noise estimator coupled to the time frequency transform logic, the background noise estimator configured to measure the continuous noise that occurs near a receiver;  
 a wind noise detector coupled to the background noise estimator, the wind detector configured to fit a line to a portion of an input signal; and  
 a wind attenuator coupled to the wind noise detector means; the wind attenuator being configured to remove a noise associated with wind that is sensed by the receiver.
- 24.** A method of removing a wind buffet from an input signal comprising:
- converting a time varying signal to a complex spectrum;  
 estimating a background noise;  
 detecting a wind buffet when a high correlation exists between a line and a portion of an input signal; and  
 dampening the wind buffet from the input signal.
- 25.** The method of claim 24 where the act of estimating the background noise comprises estimating the background noise when a transient is not detected.
- 26.** The method of claim 24 where the act of removing the wind buffet signal comprises substantially removing the wind buffet from the input signal.
- 27.** A method of removing a wind buffet from an input signal comprising:
- converting a time varying signal to a complex spectrum;  
 estimating a background noise;  
 detecting a wind buffet when a high correlation exists between a line and a portion of an input signal; and  
 removing the wind buffet from the input signal.
- 28.** A signal-bearing medium having software that controls a detection of a noise associated with a wind comprising:
- a detector that converts sound waves into electrical signals;  
 a spectral conversion logic that converts the electrical signals from a first domain to a second domain; and  
 a signal analysis logic that models a portion of the sound waves that are associated with the wind.
- 29.** The signal-bearing medium of claim 28 further comprising logic that derives a portion of a voiced signal masked by the noise.
- 30.** The signal-bearing medium of claim 28 further comprising logic that attenuates portion of the sound waves.
- 31.** The signal-bearing medium of claim 28 further comprising attenuator logic operable to limit a power in a low frequency range.
- 32.** The signal-bearing medium of claim 28 further comprising noise estimation logic that measures a continuous or ambient noise sensed by the detector.
- 33.** The signal-bearing medium of claim 32 further comprising transient logic that disables the estimation logic when an increase in power is detected.
- 34.** The signal-bearing medium of claim 28 where the signal analysis logic is coupled to an audio system.
- 35.** The signal-bearing medium of claim 28 where the signal analysis logic models only the sound waves that are associated with the wind.

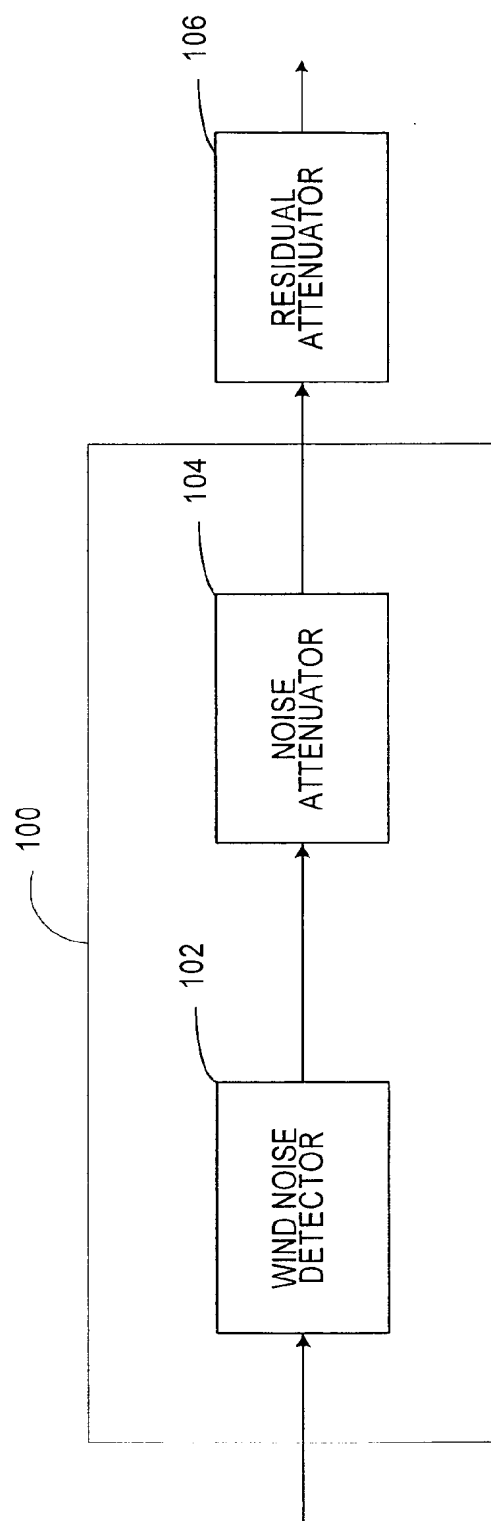


FIGURE 1

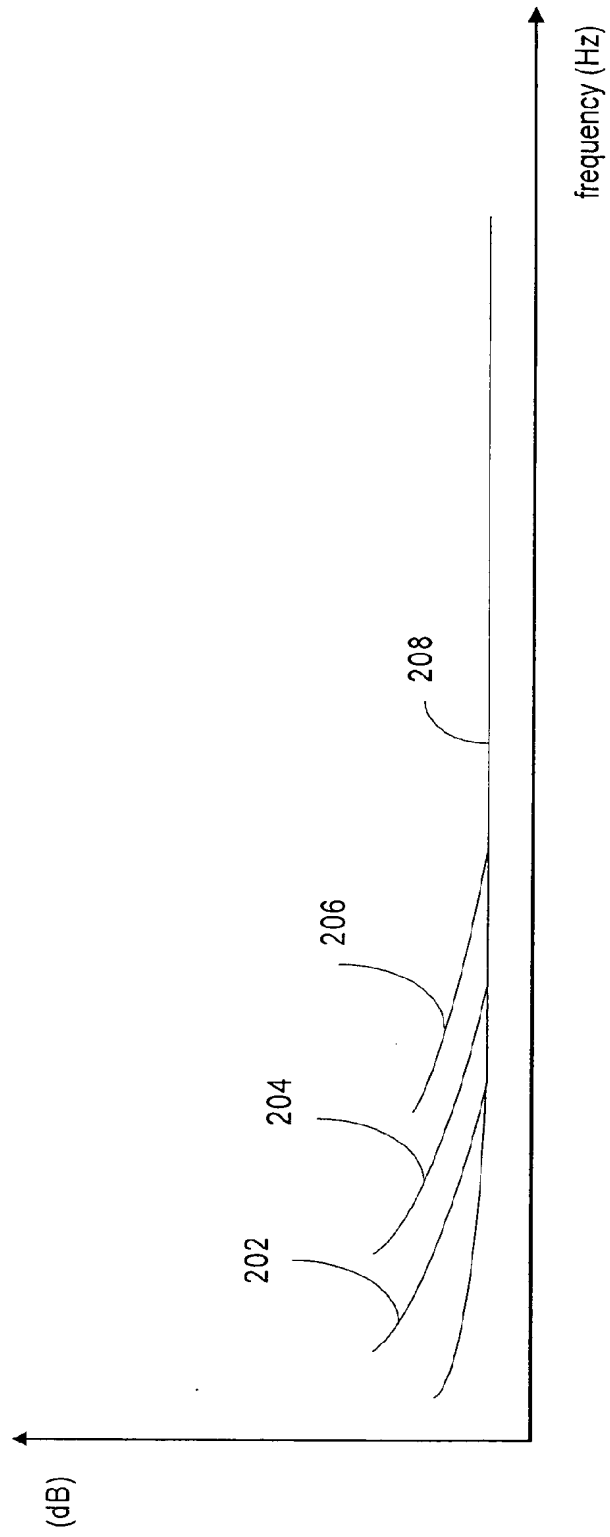


FIGURE 2

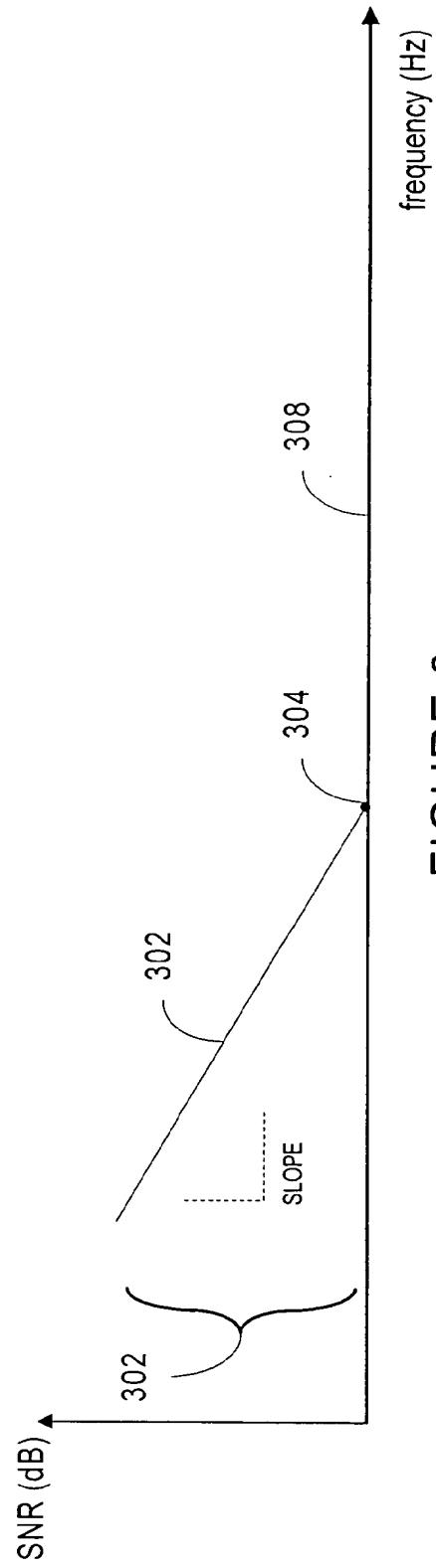


FIGURE 3

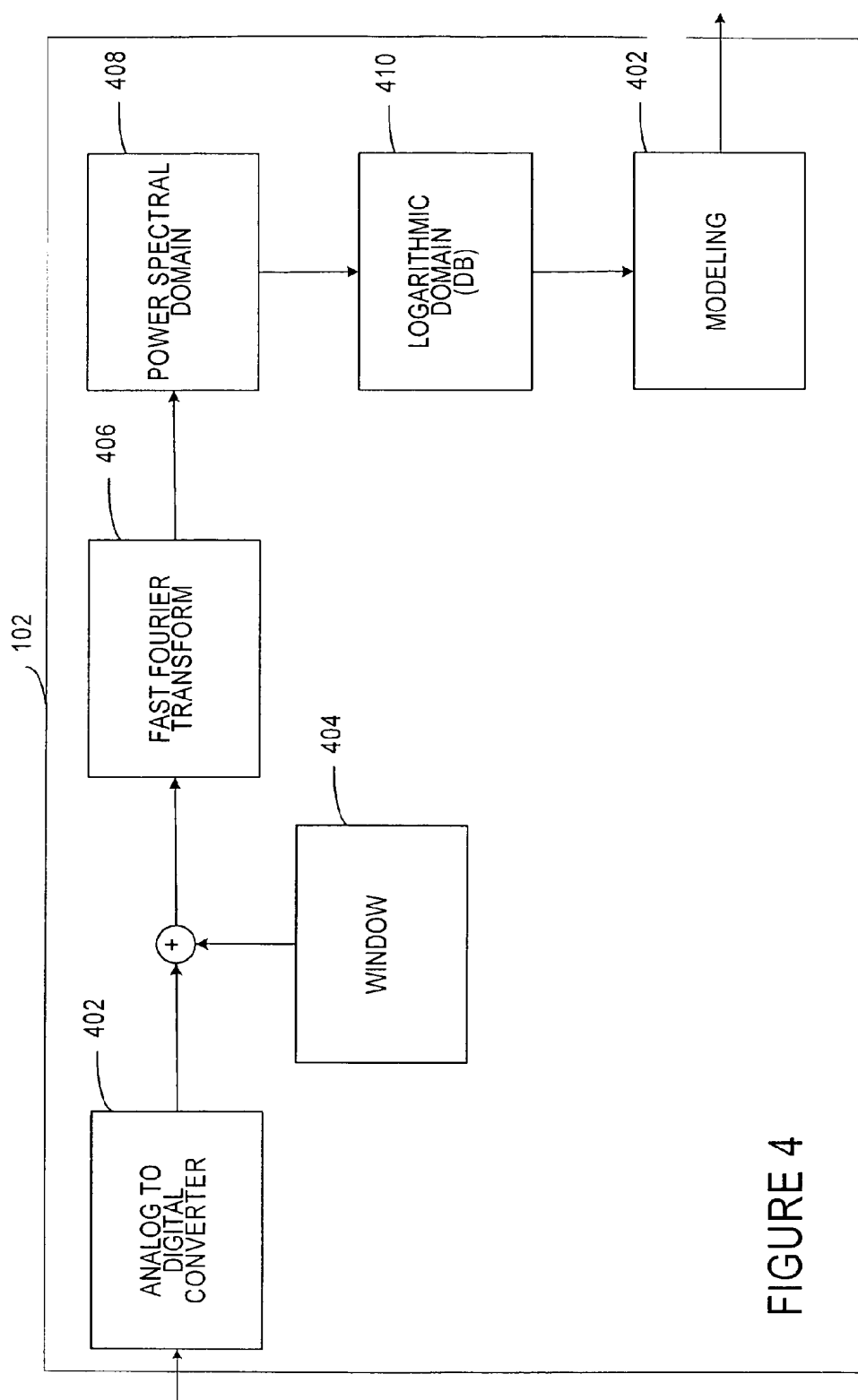


FIGURE 4

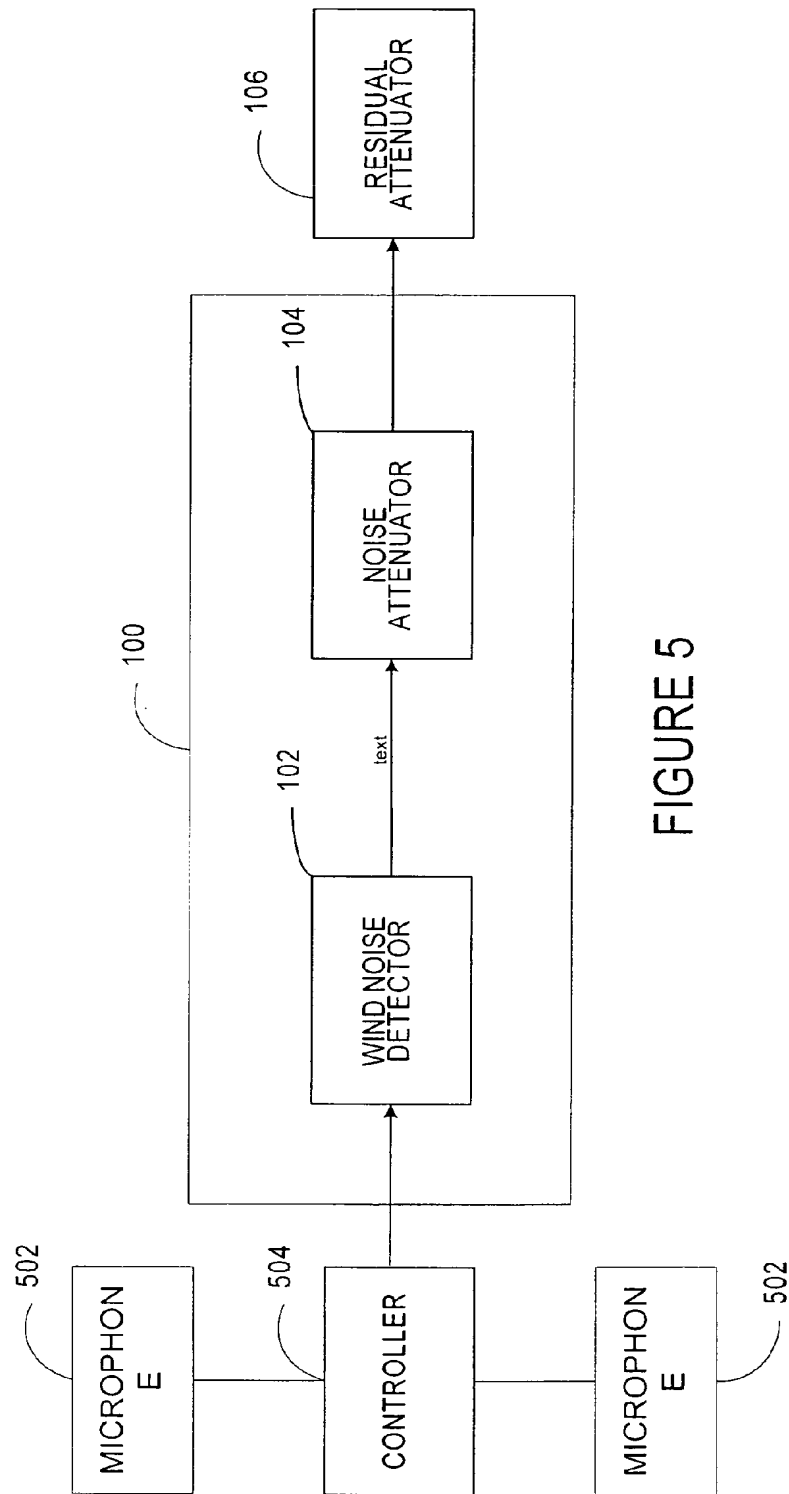


FIGURE 5

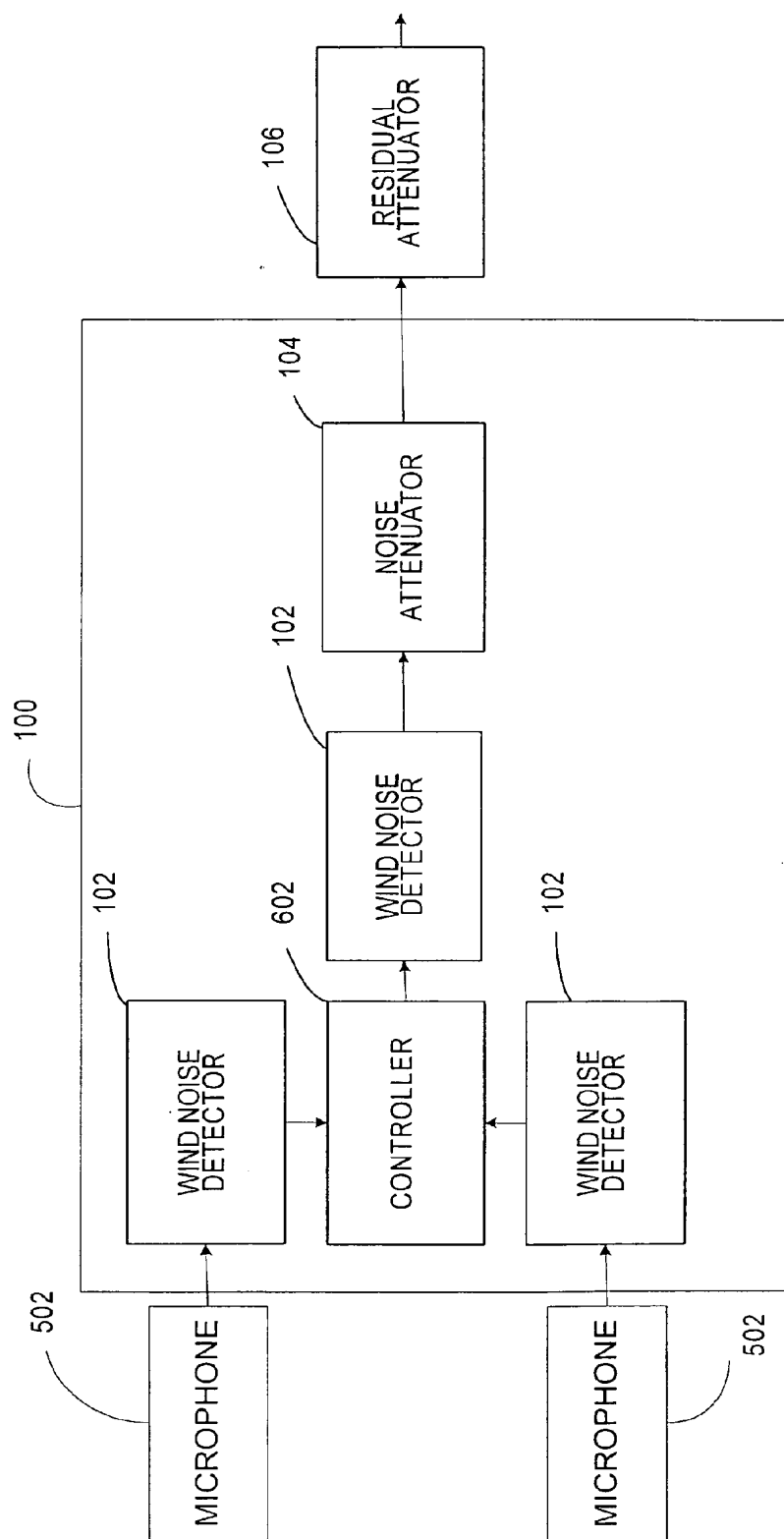


FIGURE 6

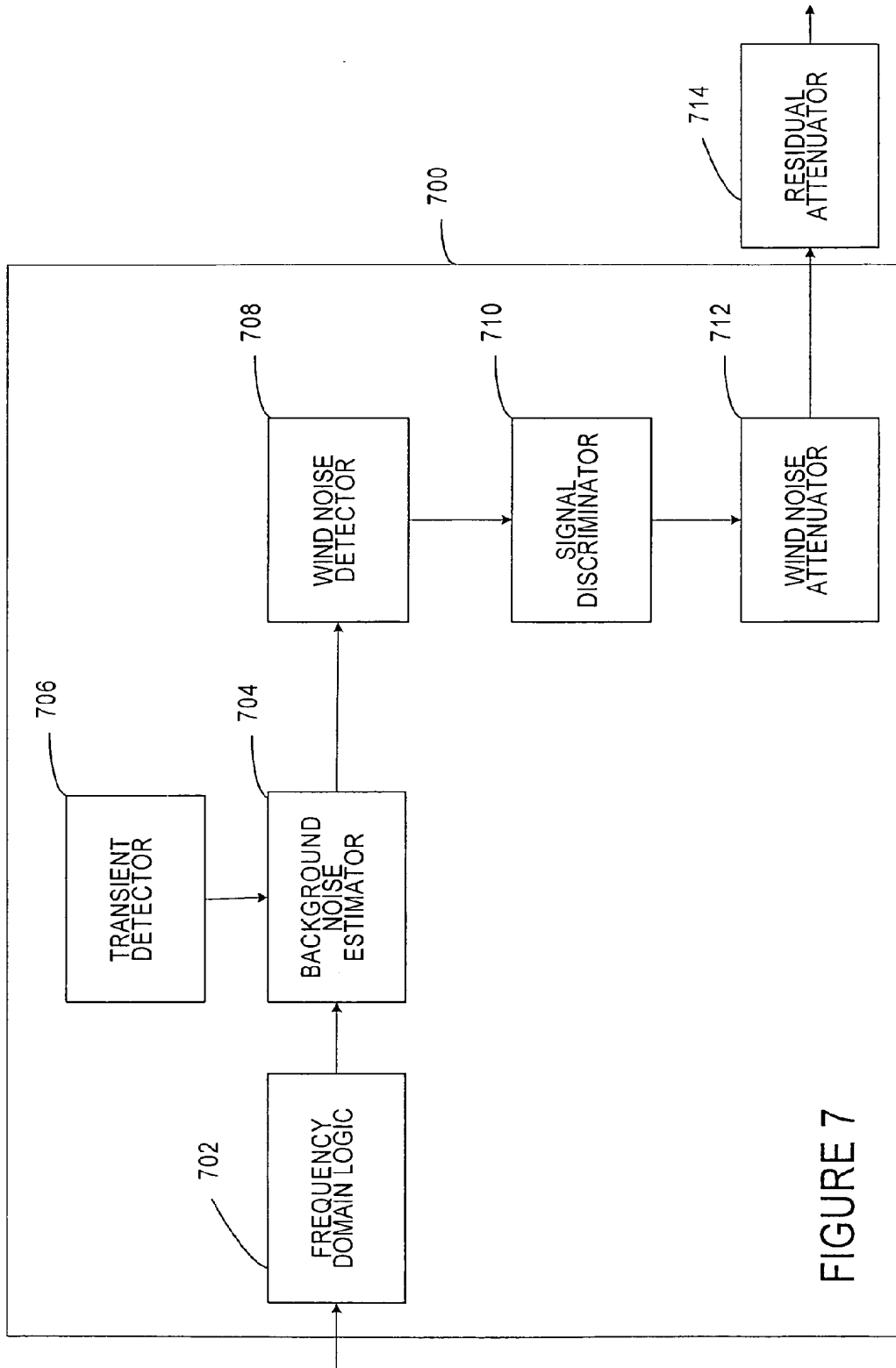


FIGURE 7



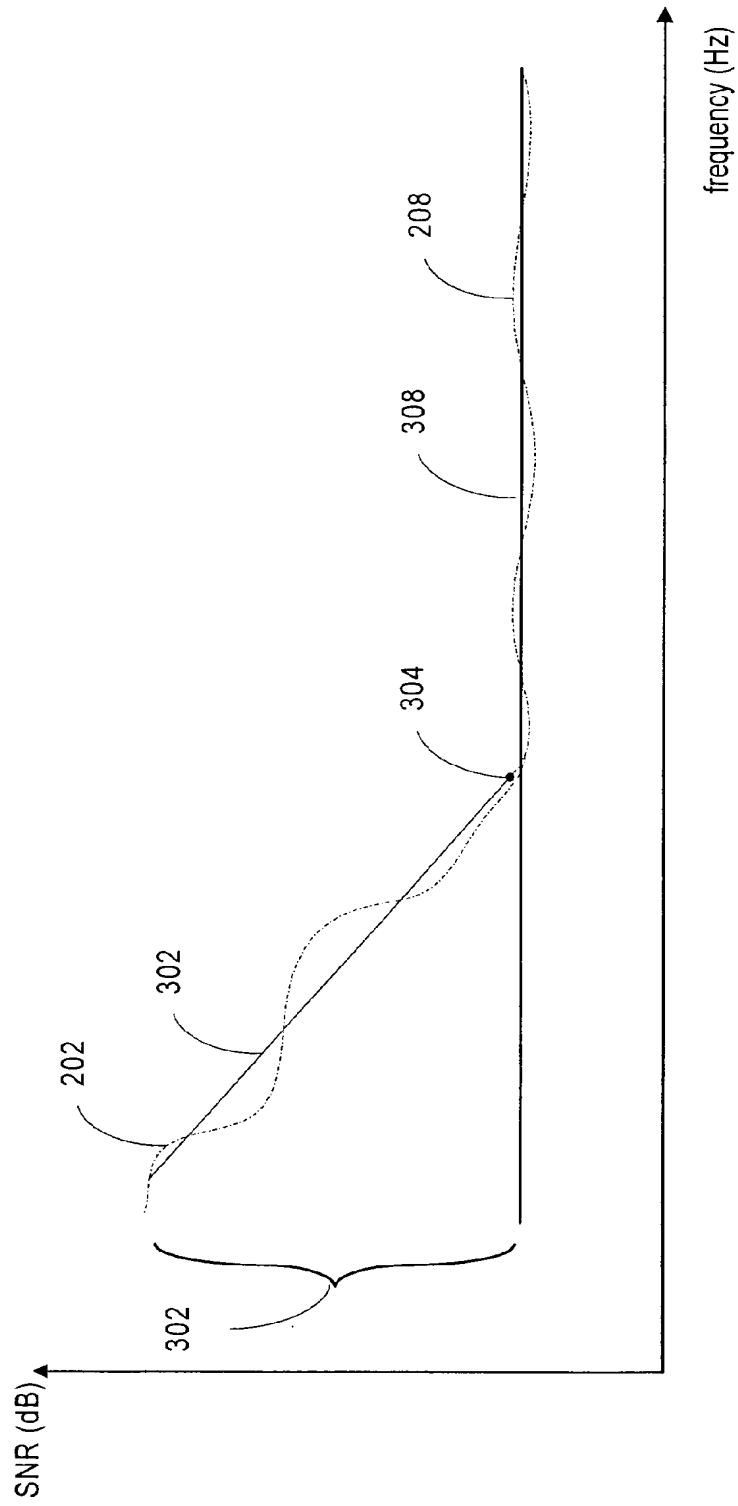


FIGURE 8

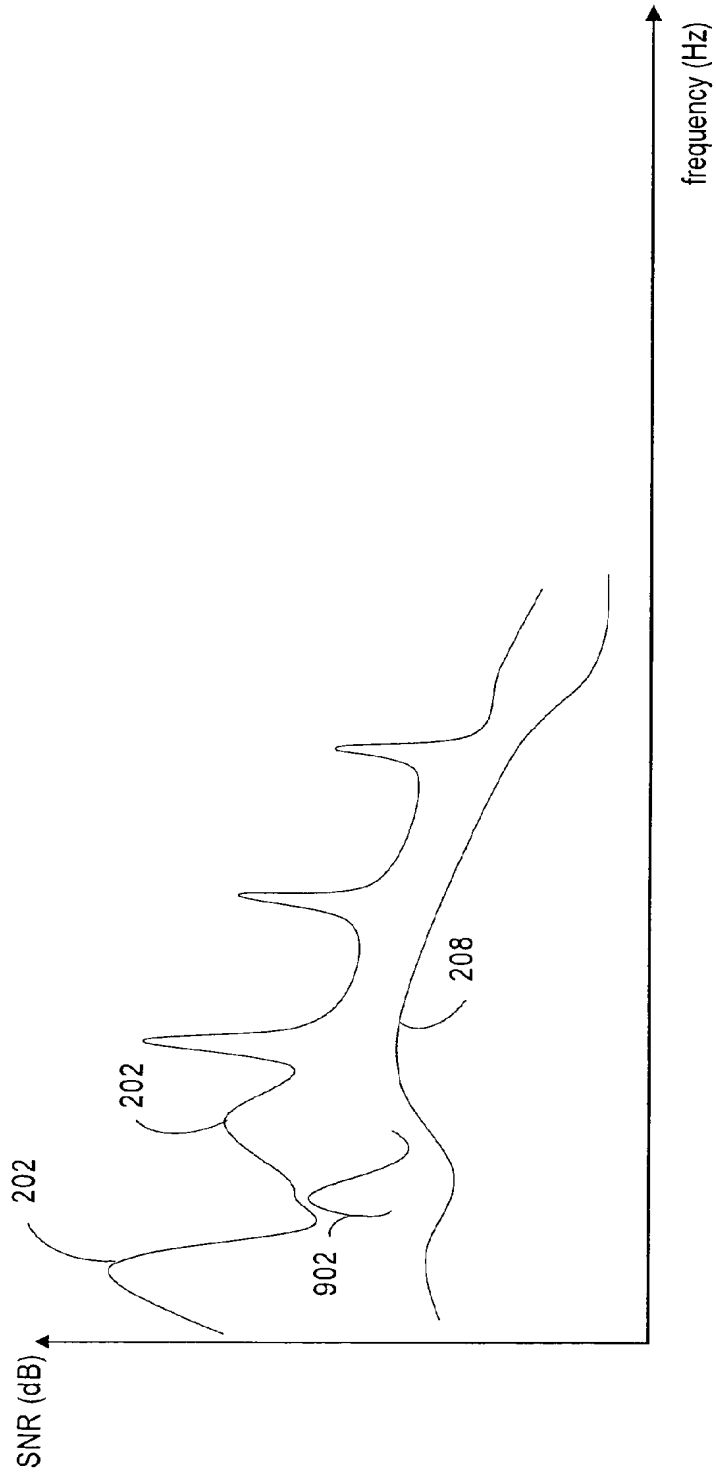


FIGURE 9

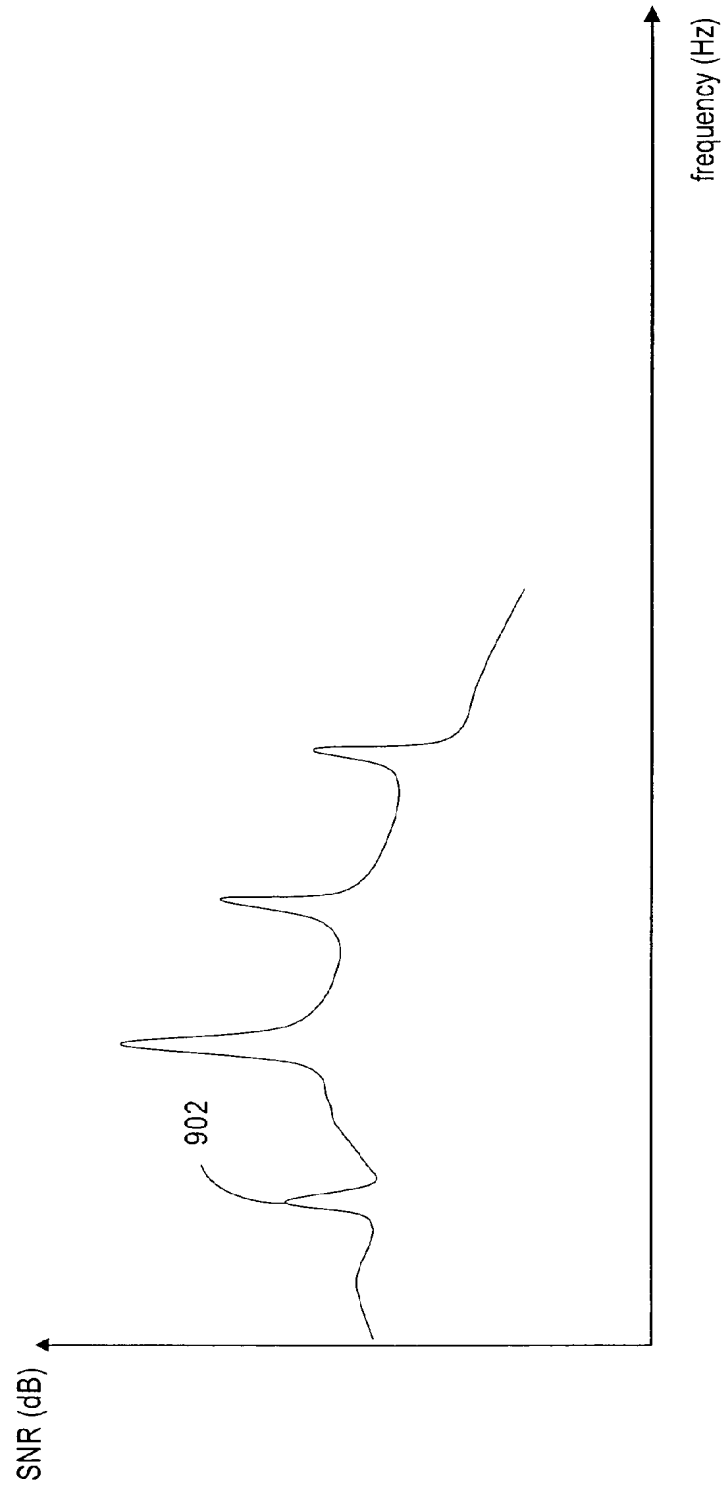


FIGURE 10

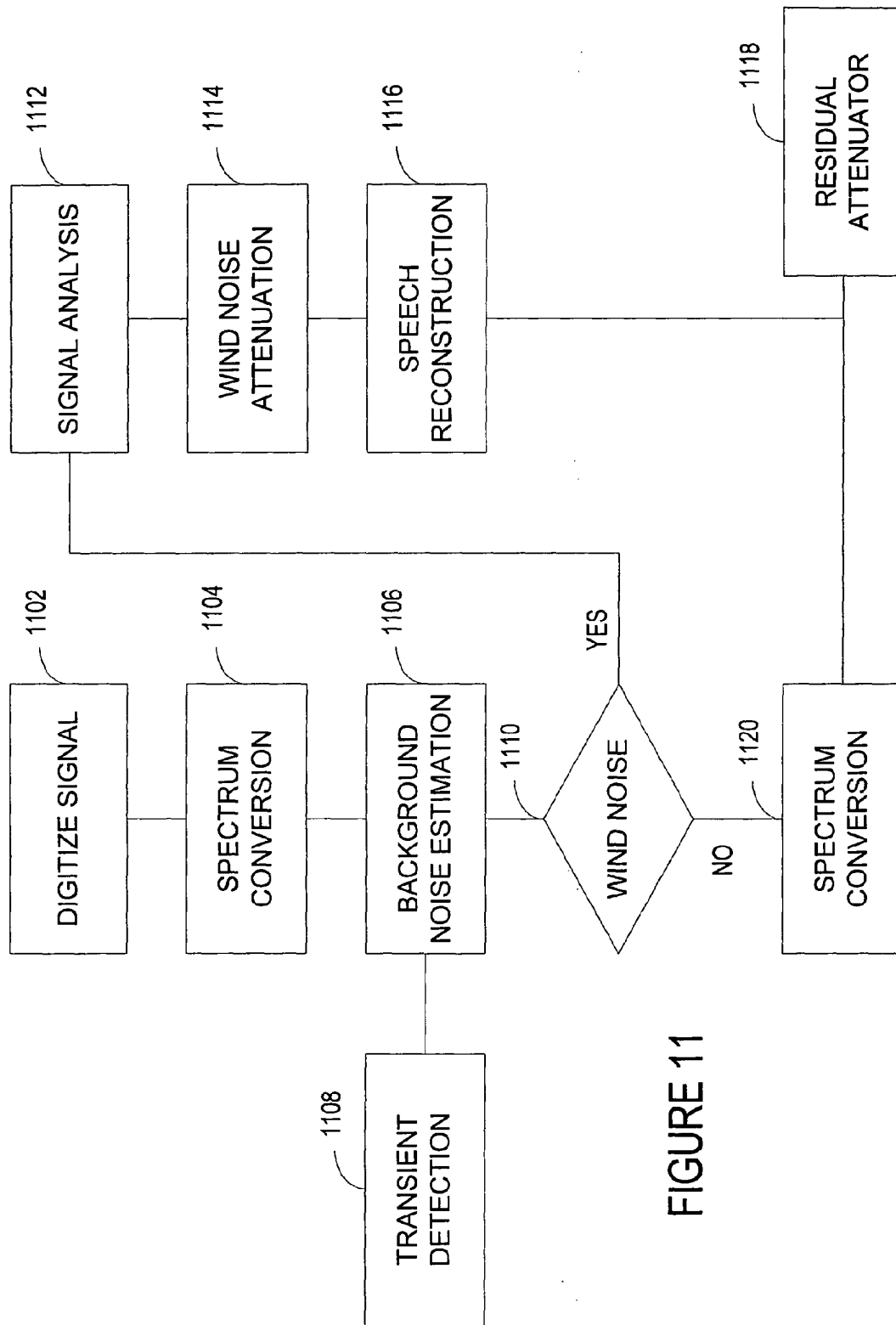


FIGURE 11

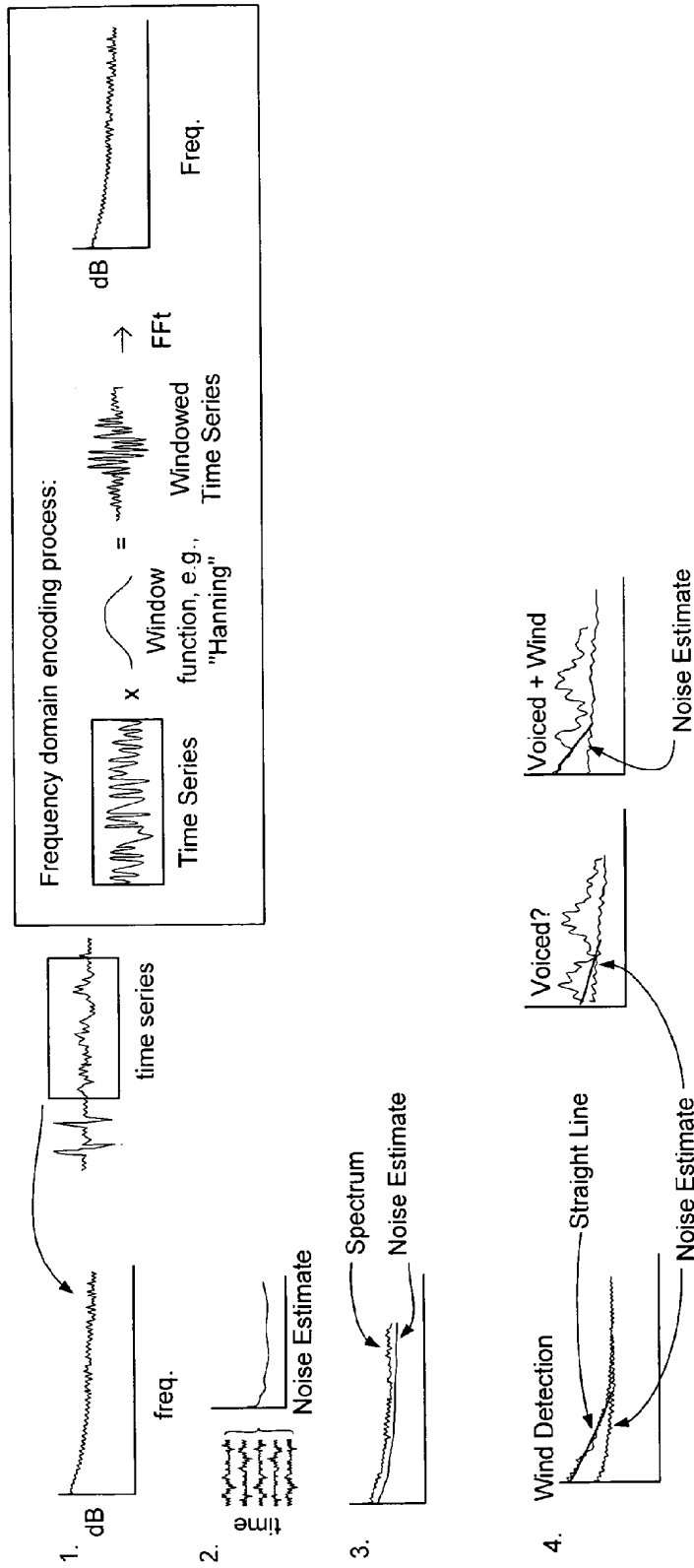


FIGURE 12

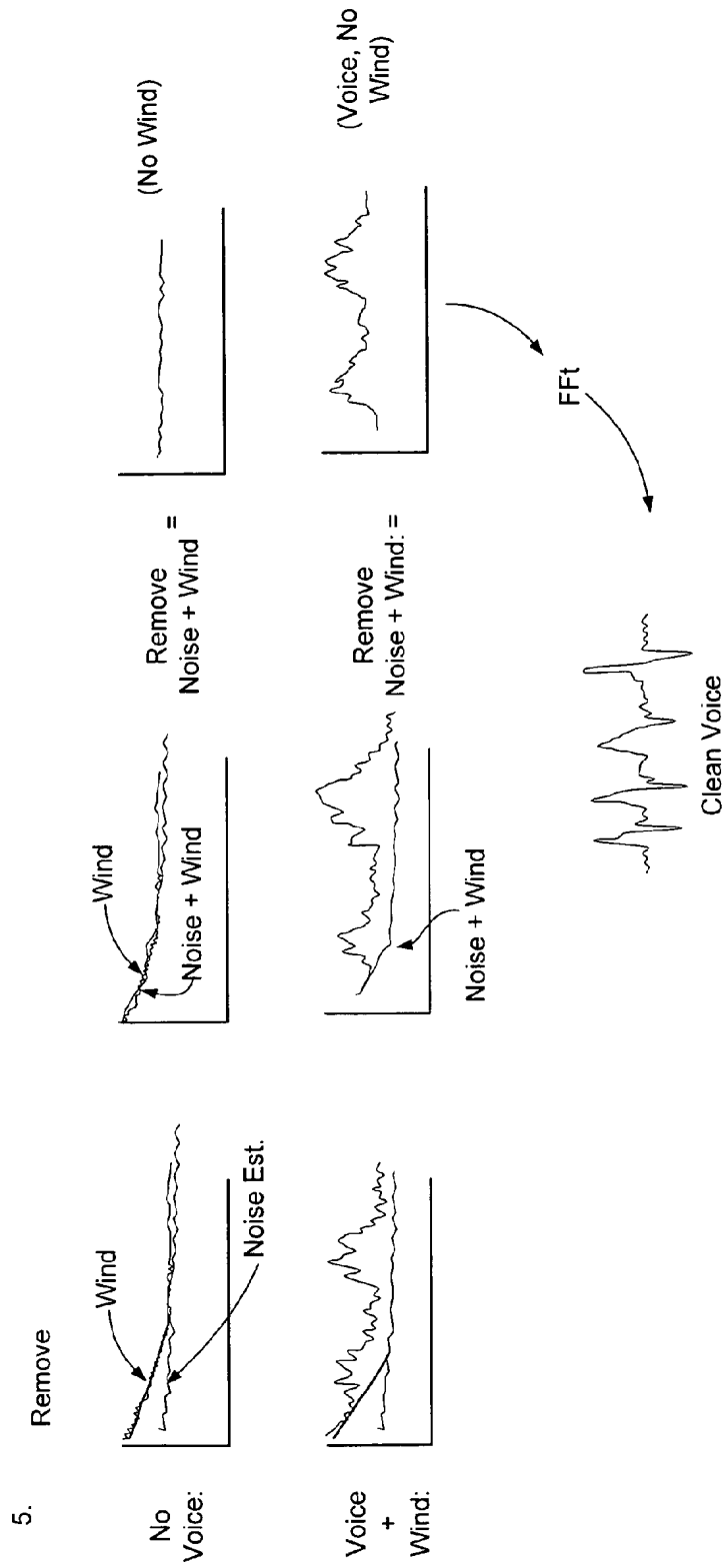


FIGURE 13

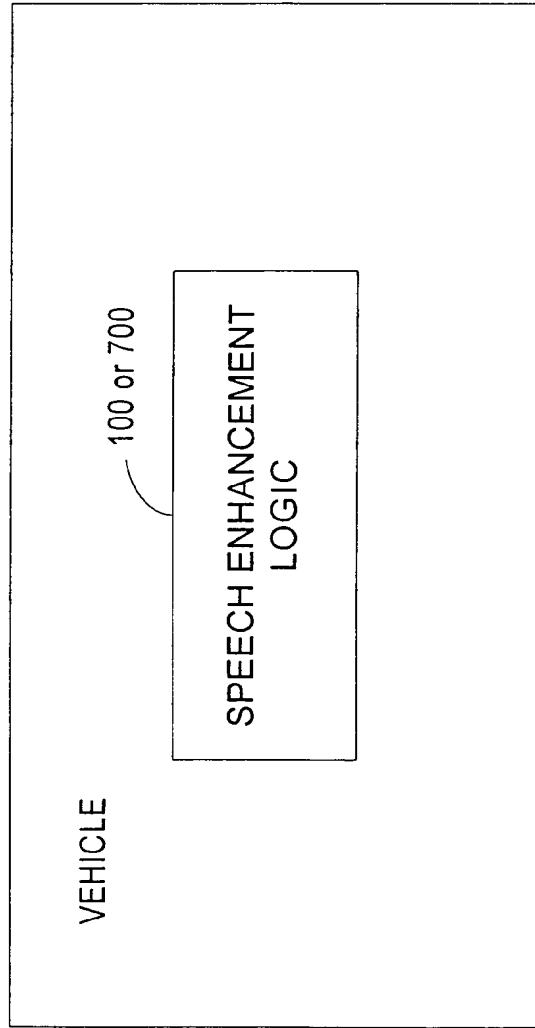


FIGURE 14

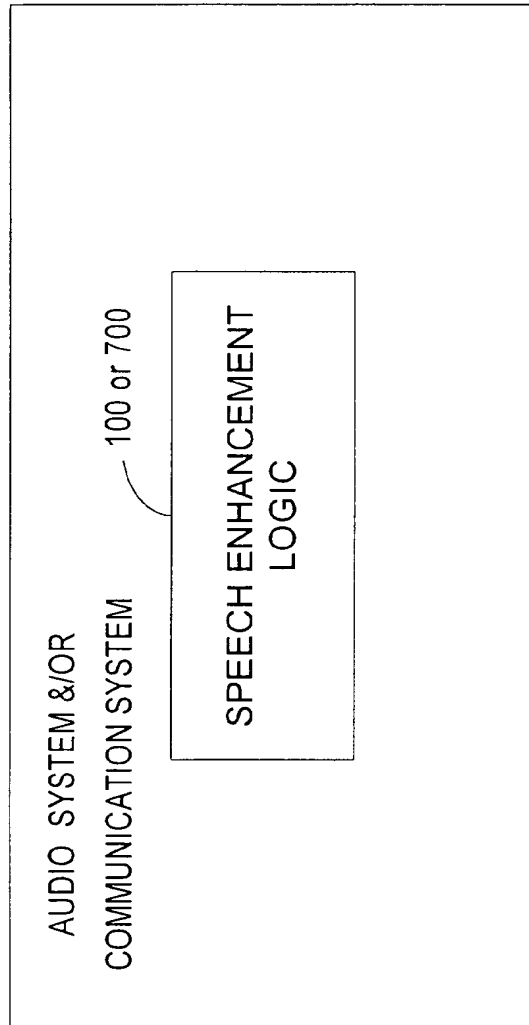


FIGURE 15





European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number  
EP 04 00 3675

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	PATENT ABSTRACTS OF JAPAN vol. 018, no. 681 (E-1649), 21 December 1994 (1994-12-21) & JP 06 269084 A (SONY CORP), 22 September 1994 (1994-09-22)	1,4,5,8, 9,11-14	G10L21/02
Y	* abstract; figure 1 *	6,7,10, 15	
A	---	2,3	
X	PUDER H ET AL: "Improved noise reduction for hands-free car phones utilizing information on vehicle and engine speeds" SIGNAL PROCESSING X THEORIES AND APPLICATIONS. PROCEEDINGS OF EUSIPCO 2000. TENTH EUROPEAN SIGNAL PROCESSING CONFERENCE, PROCEEDINGS OF 10TH EUROPEAN SIGNAL PROCESSING CONFERENCE, TAMPERE, FINLAND, 4-8 SEPT. 2000, pages 1851-1854 vol.3, XP009030255 2000, Tampere, Finland, Tampere Univ. Technology, Finland ISBN: 952-15-0443-9	16,18, 20,21, 23,24, 26-30,32	
Y	* abstract *	6,7,15, 17,19, 22,25, 31,33	G10L
	* paragraph [02.2] * * paragraph [0003] - paragraph [03.2] *		
Y	US 2001/028713 A1 (WALKER MICHAEL) 11 October 2001 (2001-10-11) * abstract; claim 1; figure 2 *	17,19, 25,33	
Y	US 2002/071573 A1 (FINN BRIAN M) 13 June 2002 (2002-06-13) * paragraphs [0071],[0085] *	10,22,31	
	---		
	-/--		
The present search report has been drawn up for all claims			
Place of search MUNICH		Date of completion of the search 30 April 2004	Examiner De Vos, L
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03.82 (p04c01)



European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number  
EP 04 00 3675

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
A	WAHAB A ET AL: "Intelligent dashboard with speech enhancement" INFORMATION, COMMUNICATIONS AND SIGNAL PROCESSING, 1997. ICICS., PROCEEDINGS OF 1997 INTERNATIONAL CONFERENCE ON SINGAPORE 9-12 SEPT. 1997, NEW YORK, NY, USA, IEEE, US, 9 September 1997 (1997-09-09), pages 993-997, XP010263974 ISBN: 0-7803-3676-3 * abstract; figures 1,2 * * paragraphs [0002],[0003] * -----	1-35	
			TECHNICAL FIELDS SEARCHED (Int.Cl.7)
The present search report has been drawn up for all claims			
Place of search MUNICH		Date of completion of the search 30 April 2004	Examiner De Vos, L
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03.82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.**

EP 04 00 3675

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.  
The members are as contained in the European Patent Office EDP file on  
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

30-04-2004

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
JP 06269084 A	22-09-1994	JP 3186892 B2	11-07-2001
US 2001028713 A1	11-10-2001	DE 10017646 A1	11-10-2001
		AU 3336101 A	11-10-2001
		CN 1325222 A	05-12-2001
		EP 1143416 A2	10-10-2001
		HU 0101288 A2	28-12-2001
		JP 2001350498 A	21-12-2001
US 2002071573 A1	13-06-2002	US 6496581 B1	17-12-2002
		US 6295364 B1	25-09-2001
		WO 02069611 A1	06-09-2002
		CA 2242510 A1	11-03-1999
		EP 0903726 A2	24-03-1999
		AU 752528 B2	19-09-2002
		AU 2253499 A	14-10-1999
		EP 0949795 A2	13-10-1999